JOURNAL OF MULTIVARIATE ANALYSIS **63**, 259–276 (1997)
ARTICLE NO. MV971696

Compact Matrix Expressions for Generalized Wald Tests of Equality of Moment Vectors*

Albert Satorra†

Universitat Pompeu Fabra, Barcelona, Spain

and

Heinz Neudecker

Universiteit van Amsterdam, Amsterdam, The Netherlands

Asymptotic chi-squared test statistics for testing the equality of moment vectors are developed. The test statistics proposed are generalized Wald test statistics that specialize for different settings by inserting an appropriate asymptotic variance matrix of sample moments. Scaled test statistics are also considered for dealing with nonstandard conditions. The specialization will be carried out for testing the equality of multinomial populations, and the equality of variance and correlation matrices for both normal and nonnormal data. When testing the equality of correlation matrices, a scaled version of the normal theory chi-squared statistic is proven to be an asymptotically exact chi-squared statistic in the case of elliptical data. © 1997 Academic Press

1. INTRODUCTION

Testing the equality of population moments is of wide generality in multivariate analysis. Specific examples are the test of equality of variance or correlation matrices and the test of the hypothesis of equality of means and

Received August 30, 1995; revised March 24, 1997.

AMS 1990 subject classifications: 62E20, 62H10, 62H15.

Key words and phrases: chi-squared test, distribution-free, ellipticity, moment vectors, multinomial distribution, normal-theory, scaled test statistic, variance and correlation matrices, Wald test.

* Paper presented at the 50th Session of the International Statistical Institute, Beijing, 21–29 August 1995. Work supported by the Spanish DGICYT Grant PB91-0814 and PB93-0403.

† Communicating author. E-mail: satorra@upf.es.

0047-259X/97 $25.00

variances across populations. Testing the equality of moments arises also in meta-analysis studies, where the results of independent studies have to be compared (Hedges and Olkin, 1985). In the present paper a family of generalized Wald test statistics will be considered. The tests to be developed share a common general formulation that is easily adapted to different settings by just inserting an appropriate asymptotic variance matrix of sample moments. The test statistics will be specialized to the test of equality of variance or correlation matrices for the cases of normality, ellipticity, and also in distribution-free settings.

The test of equality of variance matrices is usualy carried out under the assumption that the variables are normally distributed (Anderson, 1984, Chap. 10). Under the normality assumption also, a very common test for the equality of variance matrices is Bartlett's modified likelihood ratio test (Muirhead, 1982, pp. 298–309; Korin, 1968). Tests for equality of correlation matrices have also been considered under the normality assumption by Jennrich (1970). In practice, however, data deviate often from the normality assumption and it is of interest to develop methods that are free of the normality assumption. The test statistics developed in the present paper apply also to the case of nonnormal data. Recently, bootstrap techniques have been introduced for testing equality of variance matrices when data can be nonnormal (Zhang and Boos, 1992). Such techniques, however, require intensive computations. In contrast, our test statistics are fairly simple to compute. A general formulation for testing equality of correlation matrices appears in Modarres and Jernigan (1992), and a robust test for comparing correlation matrices was also investigated recently by Modarres and Jernigan (1993).

The tests to be discussed are of an asymptotic nature and they require consistent estimates of variances of sample moments, estimates which may have large variances in moderate to large sample sizes. In some cases, the matrices to be inverted may be of huge dimensions, especially in the case of asymptotic distribution-free methods. A considerable reduction in the variance of estimates of variances of sample moments, and a reduction in the dimension of the matrices to be inverted, is attained when the assumption of normality is imposed. When the normality assumption is violated, the scaled test statistics discussed in the present paper can be a sensible alternative in some cases (see, e.g., Chou, Bentler, and Satorra, 1991, for a Monte Carlo evaluation of the performance of scaled test statistics in the context of covariance structure analysis).

The plan of the paper is as follows. Section 2 develops the general expressions for the test statistics. Section 3 specializes the test statistics to different testing settings.

With regard to notation, D and D^+ will denote respectively the "duplication" and "elimination" matrices for symmetry, so that $\operatorname{vec} A = D\mathrm{v}(A)$ for symmetric matrix A, where "vec" is the usual columnwise vectorization operator and $\mathrm{v}(A)$ is obtained from vec A after eliminating the duplicated elements due to the symmetry of A. It holds that $\mathrm{v}(A) = D^+ \operatorname{vec} A$, where $D^+ \equiv (D'D)^{-1} D'$ is the Moore–Penrose inverse of D (see Magnus and Neudecker, 1988). In the present paper the matrices D will be of varying orders to be determined by the context. We denote by E_{gg} the gth unit matrix of order G, by $1_G \equiv (1, ..., 1)'$ the G-dimensional column vector of ones, by $E \equiv 1_G 1'_G$ the $G \times G$ matrix of ones, and by e_g the gth unit column vector of order G (note that $E_{gg} = e_g e'_g$, the unit matrix). Further $A \geqslant 0$ indicates that A is a positive semidefinite matrix and $\mathscr{M}(A)$ denotes the column space of A. The matrix A^- will be any generalized inverse of A (i.e., satisfying $AA^-A = A$), whereas A^+ will be the Moore–Penrose inverse. We use the notation of $A_d \equiv I \times A$ with I denoting an identity matrix and $\times$ denoting the Hadamard product of matrices. The direct (Kronecker) product of matrices will be denoted by $\otimes$. We also introduce the duplication and elimination matrices for zero-axial symmetry $\tilde{D}$ and $\tilde{D}^+$, respectively, where vector $\operatorname{vec} A = \tilde{D} w(A)$ and $w(A)$ is obtained from vec A after eliminating the zero diagonal and upper triangular elements. Clearly $w(A) = \tilde{D}^+ \operatorname{vec} A$, where $\tilde{D}^+ = \frac{1}{2}\tilde{D}'$. Given a set of vectors a_i, $i = 1, ..., I$, we denote by $\operatorname{vec}[a_i \mid e = 1, ..., I]$ the column vector formed by stacking the vectors a_i one below the other. $A_n - A \xrightarrow{P} 0$ denotes that $A_n - A$ tend in probability to zero as $n \to \infty$; when A and A_n are positive semidefinite matrices, the rank of A_n is assumed not to vary with n. Convergence in distribution as sample size $n \to \infty$ will be denoted by "$\xrightarrow{L}$."

2. GENERAL TEST FOR EQUALITY OF POPULATION MOMENTS

In this section we develop general expressions for Wald test statistics for equality of moment vectors.

2.1. *General Formulation of the Test Statistic*

Let r_g, $g = 1, ..., G$, be a p-vector of sample statistics (usually sample moments) based on independent samples of size n_g from G populations or groups ($G \geqslant 2$). For each g, assume $r_g \xrightarrow{P} \rho_g$, as $n_g \to \infty$, where ρ_g is a p-vector of population parameters (population moments), and

$$\sqrt{n_g}(r_g - \rho_g) \xrightarrow{L} N(0, \Gamma_g), \tag{1}$$

as $n_g \to \infty$, where $\Gamma_g \geqslant 0$ is a $p \times p$ finite matrix. The asymptotic variance matrix Γ_g of $\sqrt{n_g}\, r_g$ will change with the specific moment vectors r_g considered (as illustrated, for example, in Section 3).

Consider the multisample vector of sample and population moments $r \equiv \mathrm{vec}[r_g \mid g = 1, ..., G]$ and $\rho \equiv \mathrm{vec}[\rho_g \mid g = 1, ..., G]$, respectively. Clearly, from (1) and the independence of the G samples, follows

$$\sqrt{n}(r - \rho) \xrightarrow{L} N(0, \Gamma), \tag{2}$$

as $n \to \infty$ with $n_g/n \to c_g > 0$, for $g = 1, ..., G$, where $n \equiv \sum_{g=1}^{G} n_g$ is the overall sample size, c_g is the sampling fraction for group g, and Γ is the block-diagonal matrix,

$$\Gamma = \sum_{g=1}^{G} c_g^{-1}(E_{gg} \otimes \Gamma_g). \tag{3}$$

In what follows, the expression $n \to \infty$ will be understood to imply also $n_g/n \to c_g > 0$, for $g = 1, ..., G$.

Often the following assumption of the equality of the Γ_g can be made.

Assumption A. It holds that $\Gamma_g = \bar{\Gamma}$, $g = 1, ..., G$, with $\bar{\Gamma}$ a $p \times p$ positive semidefinite matrix.

Under Assumption A, we have

$$\Gamma = \Lambda^{-1} \otimes \bar{\Gamma}, \tag{4}$$

where

$$\Lambda \equiv \sum_{g=1}^{G} \frac{n_g}{n} E_{gg}. \tag{5}$$

In the present paper we are concerned with the test of the following hypothesis of equality of population moments,

$$H_0 \colon \rho_g = \vartheta; \qquad g = 1, 2, ..., G, \tag{6}$$

where ϑ is an unknown p-dimensional parameter vector. Generalized Wald test statistics (Moore, 1977) will be developed to test this hypothesis.

It will be convenient to write H_0 as the following multiple-group moment structure,

$$H_0 \colon \rho = \Delta\vartheta, \tag{7}$$

where $\Delta \equiv 1_G \otimes I_p$ is a $Gp \times p$ matrix of full column rank and ϑ is a p-vector of unknown parameters.

Consider now the weighted least-squares (WLS) estimation of the moment structure (7), with weight matrix W of the general form

$$W \equiv \sum_{g=1}^{G} \frac{n_g}{n} (E_{gg} \otimes \bar{W}) = \Lambda \otimes \bar{W}, \tag{8}$$

where $\bar{W}$ is a $p \times p$ positive definite matrix. Note that $1'_G \Lambda 1_G = 1$ and, hence, $E\Lambda E = E$. This yields the WLS estimator

$$\hat{\vartheta} \equiv (\Delta' W \Delta)^{-1} \Delta' W r = (1'_G \Lambda \otimes I_p)\, r = \sum_{g=1}^{G} \frac{n_g}{n} r_g, \tag{9}$$

of the parameter vector ϑ, and the estimator

$$\begin{aligned} \hat{\rho} \equiv \Delta \hat{\vartheta} &= \Delta (\Delta' W \Delta)^{-1} \Delta' W r = (1_G \otimes I_p)(1'_G \Lambda \otimes I_p)\, r \\ &= (E\Lambda \otimes I_p)\, r = \sum_{g=1}^{G} \frac{n_g}{n} (1_G \otimes r_g) \end{aligned} \tag{10}$$

of ρ. The vector of residuals $e \equiv r - \Delta \hat{\vartheta}$ clearly satisfies

$$e = [(I_G - E\Lambda) \otimes I_p]\, r = Qr, \tag{11}$$

where $Q \equiv (I_G - E\Lambda) \otimes I_p$ is an idempotent matrix of rank $p(G-1)$. Note that $Q\Delta = 0$ and that the estimators $\hat{\vartheta}$ and $\hat{\rho}$ do not depend on the choice of $\bar{W}$.

In some instances we will need to consider the case where the null hypothesis H_0 holds only approximately. The alternative hypothesis will then be considered,

$$H_1 : \rho = \Delta \vartheta + n^{-1/2} \eta, \tag{12}$$

where η is a pG-vector. This is an alternative hypothesis of a sequence of local alternatives that is typically used to investigate the asymptotic nonnull distribution of test statistics (e.g., Foutz and Srivastava, 1977). Clearly, $\eta = 0$ under H_0.

Using (2), we obtain

$$\sqrt{n}\, Q(r - \rho) \xrightarrow{L} N(0, Q\Gamma Q'). \tag{13}$$

Thus, under H_1, we have the distributional result

$$\sqrt{n}\, e \xrightarrow{L} N(Q\eta, Q\Gamma Q') \tag{14}$$

when $n \to \infty$, as $Q\Delta = 0$.

Let Y be a positive semidefinite matrix of the same rank as Γ such that $Y - \Gamma \xrightarrow{P} 0$ as $n \to \infty$.

Consider the generalized Wald test statistic (Moore, 1977)

$$T \equiv ne'(QYQ')^{+} e.$$

The following lemma holds.

LEMMA 1. *Under the definitions given above,*

$$T \equiv ne'(QYQ')^{+} e = nr'Q'(QYQ')^{+} Qr; \tag{15}$$

when in addition $r \in \mathcal{M}(Y)$, *then*

$$T = nr'Q'(QYQ')^{-} Qr. \tag{16}$$

Proof. We used (11) to derive the right-hand side of (15) (note that $Q'(QYQ')^{-} Q$ is a g-inverse of QYQ'). To obtain (16), note that

$$\begin{aligned} nr'Q'(QYQ')^{-} Qr &= n\ell' YQ'(QYQ')^{-} QY\ell \\ &= n\ell' Y^{1/2} Y^{1/2} Q'(QYQ')^{-} QY^{1/2} Y^{1/2}\ell, \end{aligned}$$

where $r = Y\ell$ and $Z'(ZZ')^{+} Z = Z'(ZZ')^{-} Z$ for any matrix Z, as $A = Z'(ZZ')^{+} Z - Z'(ZZ')^{-} Z = Z'\{(ZZ')^{+} - (ZZ')^{-}\} Z$, and $\sum_{ij} a_{ij}^2 = \operatorname{trace} AA' = 0$ implies $A = 0$. ∎

Given the distributional result (14), the following theorem is obtained.

THEOREM 1. *When* H_1 *holds,* T *of* (15) *satisfies*

$$T \xrightarrow{L} \chi^2(k, \lambda), \tag{17}$$

where $k = \operatorname{rank}(Q\Gamma Q')$ *and*

$$\lambda = \eta' Q'(Q\Gamma Q')^{+} Q\eta. \tag{18}$$

When Γ *is nonsingular, then* $k = p(G-1)$. *When* $r \in \mathcal{M}(Y)$ *and* $\eta \in \mathcal{M}(\Gamma)$ *then the Moore–Penrose inverse in* (18) *can be replaced by a generalized inverse,* T *and* λ *being invariant with respect to the choice of generalized inverse. Under* H_0, $\lambda = 0$, *since* $\eta = 0$.

Proof. Consider the spectral decomposition $Q\Gamma Q' = CUC'$, where $C'C = I_k$, and $U = U_d > 0$. By premultiplying both sides of (14) by C', we obtain

$$\sqrt{n}\, C'e \xrightarrow{L} N(C'Q\eta, U), \tag{19}$$

since $C'Q\Gamma Q'C = U$. Hence,

$$ne'(Q\Gamma Q')^{+}\, e = ne'CU^{-1}C'e \xrightarrow{L} \chi^2(k, \lambda).$$

The proof concludes by noting that $(QYQ')^{+} - (Q\Gamma Q')^{+} \xrightarrow{P} 0$ since $Y - \Gamma \xrightarrow{P} 0$ and generally $\operatorname{plim} A^{+} = (\operatorname{plim} A)^{+}$ from the four defining equations for the Moore–Penrose inverse (here "plim" denotes limit in probability as $n \to \infty$.) ▮

Note that the stated theorem would hold also in the case of Q being a stochastic matrix that converges to a finite probability limit $\tilde{Q}$.[1] See also Andrews (1987) for some key remarks concerning conditions for the construction of generalized Wald test statistics in the case of singular variance matrices.

2.2. *Alternative Expressions of the Test Statistic*

Some alternative expressions for the test statistic T will now be developed.

Since Γ is a block-diagonal matrix, we take

$$Y \equiv \sum_{g=1}^{G} \frac{n_g}{n} (E_{gg} \otimes Y_g), \tag{20}$$

where the $Y_g \geqslant 0$ are $p \times p$ positive semidefinite matrices. When Assumption A holds, then

$$Q\Gamma Q' = (\Lambda^{-1} - E) \otimes \bar{\Gamma}.$$

Inspired by Assumption A the matrix Y will be taken to be of the form

$$Y = \Lambda^{-1} \otimes \bar{Y}, \tag{21}$$

where $\bar{Y} \geqslant 0$ and $\bar{Y} - \bar{\Gamma} \xrightarrow{P} 0$.

If we partition Q as

$$Q = [Q_1, Q_2, ..., Q_G], \tag{22}$$

comformably with the partition of Y, the test statistic of (15) will have the alternative expression

$$T = n\left(\sum_{g=1}^{G} Q_g r_g\right)' \left(\sum_{g=1}^{G} \frac{n_g}{n} Q_g Y_g Q_g'\right)^{+} \left(\sum_{g=1}^{G} Q_g r_g\right). \tag{23}$$

The following two lemmas will be of use.

[1] Suppose $Q - \tilde{Q} \xrightarrow{P} 0$, then $k = \operatorname{rank}(\tilde{Q}\Gamma\tilde{Q}')$ and $\lambda = \eta'\tilde{Q}'(\tilde{Q}\Gamma\tilde{Q}')^{+}\, \tilde{Q}\eta$.

LEMMA 2. *Under the definitions given above, when Y is positive definite, we have*

$$Q'(QYQ')^{-}Q = Y^{-1} - Y^{-1}\Delta(\Delta'Y^{-1}\Delta)^{-}\Delta'Y^{-1}$$
$$= \Delta_{\perp}(\Delta'_{\perp}Y\Delta_{\perp})^{-}\Delta'_{\perp}, \tag{24}$$

where $\Delta_{\perp}$ denotes an orthogonal complement of the matrix Δ (i.e., a matrix of full column rank such that $\Delta'_{\perp}\Delta = 0$).

Proof. It is easy to see that the matrix

$$Y^{1/2}Q'(QYQ')^{-}QY^{1/2} + Y^{-1/2}\Delta(\Delta'Y^{-1}\Delta)^{-}\Delta'Y^{-1/2}$$

is symmetric idempotent of full rank (in the above expressions, g-inverses can be replaced by Moore–Penrose inverses); hence, it equals I_{pG}. This yields the first part of the lemma. Further, we have

$$Q'(QYQ')^{-}Q = Z(Z'YZ)^{-}Z', \tag{25}$$

due to the singular-value decomposition $Q = CZ'$, with $C'C = I_{p(G-1)}$ and $Q'(QYQ')^{+}Q = ZC'(CZ'YZC')^{+}CZ' = ZC'C(Z'YZ)^{+}C'CZ' = Z(Z'YZ)^{+}Z'$. Clearly, Z is an orthogonal complement of Δ, as $Q\Delta = 0$. ∎

The lemma can be adapted to the case of positive semidefinite Y.

LEMMA 3. *Under the definitions given above, when Y is singular and $\mathcal{M}(Q') \subset \mathcal{M}(Y)$, we have*

$$Q'(QYQ')^{-}Q = Y^{+} - Y^{+}\Delta(\Delta'Y^{+}\Delta)^{-}\Delta'Y^{+}$$
$$= \Delta_{\perp}(\Delta'_{\perp}Y\Delta_{\perp})^{-}\Delta'_{\perp}. \tag{26}$$

Proof. We use the spectral decomposition $Y = \bar{Z}M\bar{Z}'$, where $\bar{Z}'\bar{Z} = I_q$, q is the rank of Y and $M = M_d > 0$. By Lemma 2 we have

$$\bar{Z}'Q'(Q\bar{Z}M\bar{Z}'Q')^{-}Q\bar{Z} = M^{-1} - M^{-1}\bar{Z}'\Delta(\Delta'\bar{Z}M^{-1}\bar{Z}'\Delta)^{-}\Delta'\bar{Z}M^{-1}$$

as $Q\bar{Z}\bar{Z}' = L'Y^{+}\bar{Z}\bar{Z}' = L'Y^{+} = Q$. (We used $\mathcal{M}(Q') \subset \mathcal{M}(Y)$; hence $Q' = Y^{+}L$ for a suitable matrix L, and $Y^{+} = \bar{Z}M^{-1}\bar{Z}'$.) Hence,

$$\bar{Z}\bar{Z}'Q'(QYQ')^{-}Q\bar{Z}\bar{Z}' = Y^{+} - Y^{+}\Delta(\Delta'Y^{+}\Delta)^{-}\Delta'Y^{+}$$

or

$$Q'(QYQ')^{-}Q = Y^{+} - Y^{+}\Delta(\Delta'Y^{+}\Delta)^{-}\Delta'Y^{+}.$$

Furthermore, since $Q' = Y^+L$, as in (25) we can write

$$Q'(QYQ')^- Q = Z(Z'YZ)^- Z',$$

where now $Z = Y^+LC$. ▮

An explicit form for $\Delta'_\perp$ can easily be seen to be given by $\Delta'_\perp \equiv J' \otimes I_p$, where J' denotes the Helmert matrix of order G with the first row omitted.[2] Partitioning J' as

$$J' = ((J')_{.J}, ..., (J')_{.G}) = \begin{pmatrix} (J')_{1.} \\ \vdots \\ (J')_{(G-1).} \end{pmatrix},$$

where $(J')_{.g}$ and $(J')_{i.}$ denote respectively the gth column and ith row of J', we have

$$(J')_{i.} = \left(\frac{1}{i\sqrt{(i+1)}} 1'_i, \frac{-1}{\sqrt{(i+1)}}, 0'_{G-i-1} \right). \tag{27}$$

Note that $\Delta'_\perp = (\Delta_{\perp 1}, ..., \Delta_{\perp g}, ..., \Delta_{\perp G})$, where $\Delta_{\perp g} = (J')_{.g} \otimes I_p$.

Consequently, when Y is nonsingular, we can use Lemma 2 to write the test statistic T of (15) as

$$T = n\left(\sum_{g=1}^{G} \Delta_{\perp g} r_g \right)' \left(\sum_{g=1}^{G} \frac{n_g}{n} \Delta_{\perp g} Y_g \Delta'_{\perp g} \right)^+ \left(\sum_{g=1}^{G} \Delta_{\perp g} r_g \right), \tag{28}$$

and, by virtue of Lemma 3, the same expression holds when Y is singular but $\mathcal{M}(Q') \subset \mathcal{M}(Y)$. Note that in the above expression for T the matrix to be inverted is of dimension $p(G-1) \times p(G-1)$, which is slightly less than the dimension $pG \times pG$ as encountered in (15).

We are now able to state the following theorem which provides a simple expression for the test statistic T of (15).

THEOREM 2. *When $Y = \Lambda^{-1} \otimes \bar{Y}$ and $r \in \mathcal{M}(Y)$, then T of* (15) *equals*

$$T = nr'(H \otimes \bar{Y}^-)\, r, \tag{29}$$

where $H \equiv \Lambda - \Lambda E \Lambda$, $E = 1_G 1'_G$ and Λ was defined in (5).

[2] Helmert matrices have been described in Searle (1982, p. 71).

Proof.

$$\begin{aligned}
T &= nr'[(I-\Lambda E)\otimes I][(\Lambda^{-1}-E)\otimes \bar{Y}]^{+}[(I-E\Lambda)\otimes I]\,r \\
&= nr'[(I-\Lambda E)\otimes I][(\Lambda^{-1}-E)^{+}\otimes \bar{Y}^{+}][(I-E\Lambda)\otimes I]\,r \\
&= nr'[(I-\Lambda E)(\Lambda^{-1}-E)^{+}(I-E\Lambda)\otimes \bar{Y}^{+}]\,r \\
&= nr'(H\otimes \bar{Y}^{+})\,r = n\ell' Y(H\otimes \bar{Y}^{+})\,Y\ell \\
&= n\ell(\Lambda^{-1}\otimes \bar{Y})(H\otimes \bar{Y}^{+})(\Lambda^{-1}\otimes \bar{Y})\,\ell = n\ell'(\Lambda^{-1}H\Lambda^{-1}\otimes \bar{Y}\bar{Y}^{+}\bar{Y})\,\ell \\
&= n\ell'(\Lambda^{-1}H\Lambda^{-1}\otimes \bar{Y}\bar{Y}^{-}\bar{Y})\,\ell = n\ell'(\Lambda^{-1}\otimes \bar{Y})(H\otimes \bar{Y}^{-})(\Lambda^{-1}\otimes \bar{Y})\,\ell \\
&= n\ell' Y(H\otimes \bar{Y}^{-})\,Y\ell = nr'(H\otimes \bar{Y}^{-})\,r,
\end{aligned}$$

where $r = Y\ell$ and the equality

$$\begin{aligned}
H &= \Lambda - \Lambda E\Lambda = \Lambda(\Lambda^{-1}-E)\,\Lambda \\
&= \Lambda(\Lambda^{-1}-E)(\Lambda^{-1}-E)^{+}(\Lambda^{-1}-E)\,\Lambda \\
&= (I-\Lambda E)(\Lambda^{-1}-E)^{+}(I-E\Lambda)
\end{aligned}$$

was used.[3] We note the properties of $H \geqslant 0$, $\operatorname{rank}(H) = G-1$ and $H1_G = 0$. ∎

In some applications one may be using a misspecified expression for $\bar{Y}$, i.e. a test statistic of the general form

$$T_{\bar{V}} \equiv nr'(H\otimes \bar{V}^{+})\,r, \tag{30}$$

where $\bar{V}-\bar{\Omega}\xrightarrow{P} 0$ with $\bar{\Omega}\geqslant 0$ a $p\times p$ matrix with possibly $\bar{\Omega}\neq\bar{\Gamma}$. We define $V=\Lambda^{-1}\otimes\bar{V}$ and $\Omega=\Lambda^{-1}\otimes\bar{\Omega}$. Note that $V-\Omega\xrightarrow{P} 0$. Following the line of proof of Theorem 2, it can easily be seen that when $r\in\mathscr{M}(V)$ then

$$T_{\bar{V}} \equiv nr'(H\otimes \bar{V}^{-})\,r. \tag{31}$$

The following theorem establishes the limit distribution of this general class of test statistics.

THEOREM 3. *Let $T_{\bar{V}}$ denote the quadratic form statistic given in* (30). *Then, under H_1,*

(*a*) $T_{\bar{V}}\xrightarrow{L}\sum_{i=1}^{k}\alpha_i(u_i+\omega_i)^2$, *where the u_i's are independent standard normal variables, the ω_i's are the components of* $\omega\equiv\Lambda^{-1}R'\Gamma^{1/2}(H\otimes\bar{\Omega}^{+})\,\eta$,

[3] We are indebted to Anna Cuxart of Universitat Pompeu Fabra for providing the proof of this equality involving a g-inverse instead of the Moore–Penrose inverse.

$\Gamma^{1/2}(H \otimes \bar{\Omega}^{+})\,\Gamma^{1/2} = RAR'$, $R'R = I_k$, $A = A_d$, *the* α_i*'s are the diagonal (positive) elements of* A, *and* $k \equiv \operatorname{rank}\{\Gamma^{1/2}(H \otimes \bar{\Omega}^{+})\Gamma^{1/2}\}$;

(*b*) *when* $\Gamma = \Lambda^{-1} \otimes \bar{\Gamma}$ *and* $\bar{\Omega} = \bar{\Gamma}$ *then* $T_{\bar{V}}$ *is the statistic* T *of* (15) *and thus the result* (17) *applies.*

Note that under H_0 *we have* $\eta = 0$, *and hence,* $\omega = 0$. *When* $r \in \mathscr{M}(V)$ *then the Moore–Penrose inverse can be replaced by a generalized inverse.*

Proof. Note that

$$T_{\bar{V}} \equiv nr'(H \otimes \bar{V}^{+})\,r = n(r - \Delta\vartheta)'\,(H \otimes \bar{V}^{+})(r - \Delta\vartheta), \tag{32}$$

since $H1_G = 0$. Consequently, since under H_1

$$\sqrt{n}(r - \Delta\vartheta) \xrightarrow{L} \mathscr{N}(\eta, \Gamma),$$

we obtain that

$$T_{\bar{V}} \xrightarrow{L} z'(H \otimes \bar{\Omega}^{+})\,z, \tag{33}$$

where $z = \mathscr{N}(\eta, \Gamma)$. The theorem now follows by straightforward application of standard results on quadratic forms in normal variables (see Dik and Gunst, 1985; also Neudecker, 1994), applied to the right-hand side of (33). ∎

2.3. *Scaled Chi-Squared Test Statistics*

As in Rao and Scott (1984), when the asymptotic chi-squaredness of the test statistic is not guaranteed, it may be of interest to consider a first-order adjustment of the not necessarily asymptotic chi-squared test statistic. Consider the scaled statistic

$$\bar{T}_{\bar{V}} = T_{\bar{V}}/a, \tag{34}$$

where $T_{\bar{V}}$ is defined in (30) and a is a consistent estimator of

$$\alpha \equiv \frac{1}{k} \operatorname{tr} Q'(Q\Omega Q')^{+}\,Q\Gamma,$$

where k is given in Theorem 3. Note that Theorem 3 implies that under H_0, the asymptotic mean of $T_{\bar{V}}$ is equal to $\operatorname{tr}\{Q'(Q\Omega Q')^{+}\,Q\Gamma\}$; thus under H_0 the asymptotic mean of the scaled statistic $\bar{T}_{\bar{V}}$ is the same as the mean of the χ^2_k. This suggests the use of $\bar{T}_{\bar{V}}$ as an approximately chi-squared statistic when $T_{\bar{V}}$ is not exactly chi-squared. See Satorra and Bentler (1994) for a similar scaled test statistic in the context of covariance structure analysis.

Generally, we have

$$\alpha = \operatorname{tr}\{(Q\Omega Q')^{+}\,(Q\Gamma Q')\}/k$$
$$= \operatorname{tr}(\Theta' Q\Gamma Q'\Theta)/k = \sum_{g=1}^{G} \frac{n}{n_g} \operatorname{tr}(\Theta' Q_g \Gamma_g Q_g' \Theta)/k,$$

where $\Theta\Theta' = (Q\Omega Q')^{+}$. A consistent estimator of α can thus be constructed as

$$a = \sum_{g=1}^{G} \frac{n}{n_g} \operatorname{tr}\{B' Q_g Y_g Q_g' B\}/k, \tag{35}$$

where $BB' = (QVQ')^{+}$. Often, sample moments are of the form

$$r_g \equiv \frac{1}{n_g} \sum_{i=1}^{n_g} d_{gi}, \qquad g = 1, ..., G, \tag{36}$$

where the $\{d_{gi}\}_{i=1}^{n_g}$, $g = 1, ..., G$, are mutually independent i.i.d. (independent and identically distributed) sequences of p-dimensional random vectors. In this setup an (asymptotic distribution-free) consistent estimator of Γ_g will be

$$Y_g = \frac{1}{n_g} \sum_{i=1}^{n_g} (d_{gi} - \bar{d}_g)(d_{gi} - \bar{d}_g)', \tag{37}$$

where $\bar{d}_g$ denotes the sample mean of $\{d_{gi}\}_{i=1}^{n_g}$. Further, under Assumption A, the corresponding "pooled" estimator of the common matrix $\bar{\Gamma}$ will be $Y = \Lambda^{-1} \otimes \bar{Y}$ with

$$\bar{Y} = \sum_{g=1}^{G} \frac{n_g}{n} Y_g. \tag{38}$$

Thus in the case of i.i.d. sampling, we can write (35) as

$$a = \frac{1}{k} \sum_{g=1}^{G} \frac{n}{n_g} \left(\frac{1}{n_g} \sum_{i=1}^{n_g} b_{gi}' b_{gi} \right),$$

where $b_{gi} \equiv B' Q_g (d_{gi} - \bar{d}_g)$. Further, when $\Gamma = \Lambda^{-1} \otimes \bar{\Gamma}$ holds, then

$$Q'(Q\Omega Q')^{+}\, Q\Gamma = (H \otimes \bar{\Omega}^{+})(\Lambda^{-1} \otimes \bar{\Gamma}) = H\Lambda^{-1} \otimes \bar{\Omega}^{+}\bar{\Gamma};$$

hence

$$\alpha = \frac{G-1}{k} \operatorname{tr} \bar{\Omega}^{+}\bar{\Gamma},$$

since $\operatorname{tr} H\Lambda^{-1} = G-1$. Thus a consistent estimator of α in that case will be

$$a = \frac{G-1}{k} \operatorname{tr} \bar{V}^{+} \bar{Y}.$$

Note that the scaling correction will be automaticaly inactive (asymptotically) when $\operatorname{tr} \bar{\Omega}^{+} \bar{\Gamma} = p$ and $k = (G-1)p$. Note also that $\operatorname{tr} \bar{\Omega}^{+} \bar{\Gamma} = p$ when $\bar{\Omega}^{+} = \bar{\Gamma}^{-1}$. When also $r \in \mathcal{M}(V)$ the Moore–Penrose inverse can be replaced by a g-inverse.

In some nonstandard conditions, as for example in multistage clustered sampling, the estimator a of α would have the same expression as above, but with the expression Y_g of (37) modified so that the new Y_g is a consistent estimator of Γ_g.

3. APPLICATIONS TO SPECIFIC TESTING SETTINGS

The above described test statistics will now be applied to different cases of equality of moment matrices.

3.1. *Equality of Multinomial Populations*

Consider the problem of

$$H_0: \rho_g = \bar{\rho}, \qquad g = 1, ..., G,$$

where ρ_g is a p-dimensional vector of positive numbers (proportions) satisfying $1'_p \rho = 1$. Let r_g be the p-vector of sample proportions for which $1'_p r_g = 1$, $g = 1, ..., G$, and r_g converges to ρ_g in probability as sample size $n_g \to \infty$. In the case of a multinomial distribution the variance matrix of r_g is known to be

$$\Gamma_g = P_g - \rho_g \rho'_g, \tag{39}$$

where $P_g = \operatorname{dg}(\rho_g)$. Here $\operatorname{dg}(a)$ for a vector a denotes the diagonal matrix with the elements of a on the diagonal. Note that Γ_g is of rank $p-1$. We define the pooled estimator of the common matrix Γ_g as

$$\bar{Y} \equiv \bar{R} - \bar{r}\bar{r}',$$

where $\bar{r} = \sum_{g=1}^{G} (n_g/n)\, r_g$ and $\bar{R} = \operatorname{dg}(\bar{r})$, and we let $Y \equiv \Lambda^{-1} \otimes \bar{Y}$. We have that under H_0, $Y - \Gamma \xrightarrow{P} 0$. Note that in the present testing setting, the null hypothesis H_0 implies Assumption A of equality of the variance matrices Γ_g.

Let $r \equiv \text{vec}[r_g \mid g=1, ..., G]$. Note that $(r-(1/p)\,1_{pG}) \in \mathcal{M}(Y)$, since $\bar{Y}$ is of rank $p-1$, $1'_p\,\bar{Y}=0$ and $1'_p r_g = 1$. Thus, the test statistic of (29) will be

$$T = nr'(H \otimes \bar{R}^{-1})\, r, \tag{40}$$

as $nr'(H \otimes \bar{R}^{-1})\, r = n(r-(1/p)\,1_{pG})'\,(H \otimes \bar{R}^{-1})(r-(1/p)\,1_{pG})$ (we use again $1'_G H = 0$) and $\bar{R}^{-1}$ is generalized inverse of the variance matrix $\bar{Y} = \bar{R} - \bar{r}\bar{r}'$. Further, it is easy to see that $k = \text{rank}(Q\Gamma Q') = (G-1)(p-1)$, since $\text{rank}(\bar{\Gamma}) = p-1$.[4]

In the case where there is deviation from iid sampling of a multinomial distribution, as for example, with cluster sampling, the form of (39) is violated and the scaled statistic $\bar{T} = a^{-1}T$ may be a convenient approximate chi-squared test statistic. Under Assumption A, we have $a = ((G-1)/k)\,\text{tr}\,\bar{R}\bar{Y} = \text{tr}\,\bar{R}\bar{Y}/p$, where $\bar{Y}$ is a consistent estimator of the common variance matrix of the r_g.

3.2. *Equality of Variance and Augmented Moment Matrices*

Consider $H_0\colon \Sigma_g = \bar{\Sigma}$, $g=1, ..., G$, where Σ_g is the $h \times h$ variance matrix of the gth group (population). Consider $\{z_{gi}\}_{i=1}^{n_g}$, $g=1, ..., G$, to be mutually independent i.i.d. sequences of $h \times 1$ vectors, and

$$S_g \equiv \frac{1}{n_g} \sum_{i=1}^{n_g} (z_{gi} - \bar{z}_g)(z_{gi} - \bar{z}_g)', \tag{41}$$

be the usual (biased) sample variance matrix for the gth group. Here $\bar{z}_g$ denotes the sample mean of $\{z_{gi}\}_{i=1}^{n_g}$. We define the pooled sample variance matrix $\bar{S} \equiv \sum_{g=1}^{G} (n_g/n)\, S_g$. Define $s = \text{vec}[s_g \mid g=1, ..., G]$ and $\sigma = \text{vec}[\sigma_g \mid g=1, ..., G]$, where $s_g \equiv D^+ \text{vec}\, S_g$ and $\sigma_g \equiv D^+ \text{vec}\, \Sigma_g$. Note that s and σ are pG-dimensional vectors where $p = 2^{-1}h(h+1)$.

Denote by Φ_g the asymptotic variance matrix of $\sqrt{n_g}\, s_g$, and by $\bar{\Phi}$ the same variance matrix when it is common to all groups. Clearly, regardless of the distribution of the z_i, consistent estimators of Φ_g and $\bar{\Phi}$ are respectively

$$W_g \equiv \frac{1}{n_g} \sum_{i=1}^{n_g} (v_{gi} - s_g)(v_{g_i} - s_g)' \tag{42}$$

and

$$\bar{W} \equiv \sum_{g=1}^{G} \frac{n_g}{n}\, W_g, \tag{43}$$

where $v_{gi} = D^+ \text{vec}(z_{gi} - \bar{z}_g)(z_{gi} - \bar{z}_g)'$.

[4] $Q\Gamma Q' = [(I-E\Lambda)\otimes I_p](\Lambda^{-1}\otimes\bar{\Gamma})[(I-E\Lambda)\otimes I_p] = (I-E\Lambda)\,\Lambda^{-1}(I-E\Lambda)\otimes\bar{\Gamma} = (\Lambda^{-1}-E)\otimes\bar{\Gamma} = (I-E\Lambda)\,\Lambda^{-1}\otimes\bar{\Gamma}$. Consequently, $\text{rank}(Q\Gamma Q') = \text{rank}\{(I-E\Lambda)\,\Lambda^{-1}\}\,\text{rank}(\bar{\Gamma}) = \text{rank}\{(I-E\Lambda)\}\,\text{rank}(\bar{\Gamma}) = (G-1)(p-1)$.

When $\{z_{gi}\}_{i=1}^{n_g}$ are i.i.d. normally distributed, then Φ_g takes the following normal-theory (NT) expression (e.g., Magnus and Neudecker, 1986),

$$\Phi_g^* \equiv 2D^+(\Sigma_g \otimes \Sigma_g)\, D^{+\prime}, \tag{44}$$

which is consistently estimated by

$$W_g^* \equiv 2D^+(S_g \otimes S_g)\, D^{+\prime}. \tag{45}$$

The corresponding matrices $\bar{\Phi}^*$ and $\bar{W}^*$ are obtained by replacing the matrix Σ_g in the expression of Φ_g^* by $\bar{\Sigma}$ and $\bar{S}$, respectively. Note that under NT, the null hypothesis H_0 implies also Assumption A of equality of the matrices Φ_g.

The asymptotic distribution-free (DF) and the normal-theory (NT) form of the test statistic of (29) will thus be

$$T = ns'(H \otimes \bar{W}^-)\, s \tag{46}$$

and

$$T^* = ns'(H \otimes \bar{W}^{*-})\, s, \tag{47}$$

respectively. The number of degrees of freedom of the test is equal to $k=(G-1)\,p$. Note that T^* is computationally easier to obtain than T, since T involves the inversion of a $p \times p$ matrix, while T^* requires to invert a matrix just of dimension $h \times h$.

The scaled version (34) of T^* will be

$$\bar{T} = T^*/a, \tag{48}$$

where $a = \mathrm{tr}(\bar{W}^{*-1}\bar{W})/p$.

Consider now the case where $z_{gi} = (y'_{gi}, 1)'$ is an augmented moment vector and $S_g \equiv (1/n_g)\sum_{i=1}^{n_g} z_{gi} z'_{gi} \xrightarrow{P} \Sigma_g$, as $n_g \to \infty$, where Σ_g and S_g are called the population and sample augmented moment matrices, respectively. In this case, H_0 is the hypothesis of equality of mean vector and variance matrix across groups. It holds that W_g of (42) and $\bar{W}$ of (43) still give consistent estimators of the DF expressions of Φ_g and $\bar{\Phi}$, respectively. Now, however, the NT form of Φ_g is (e.g., Magnus and Neudecker, 1986)

$$\Phi_g^* \equiv 2D^+(\Sigma_g \otimes \Sigma_g - \mu_g \mu'_g \otimes \mu_g \mu'_g)\, D^{+\prime}, \tag{49}$$

where μ_g is the probability limit of $(1/n_g)\sum_{i=1}^{n_g} z_{gi}$. It can easily be verified that $2^{-1}D'(\Sigma_g^{-1} \otimes \Sigma_g^{-1})D = \Phi_g^{*-1}$, thus the same expressions of T, T^*, and $\bar{T}$ as reported in (46), (47), and (48), respectively, hold in the case of augmented moment matrices.

3.3. *Testing the Equality of Correlation Matrices*

Consider $H_0\colon P_g = \bar{P}$, $g = 1, ..., G$, where P_g is the $(h \times h)$ correlation matrix of an h-dimensional vector. Let $R_g \equiv (S_g)_d^{-1/2} S_g (S_g)_d^{-1/2}$ be the sample correlation matrix, where S_g is the gth sample variance matrix. Let $r \equiv \text{vec}[r_g \mid g = 1, ..., G]$, where $r_g \equiv \tilde{D} \text{ vec } R_g$ of dimension $p \equiv h(h-1)/2$. The asymptotic variance matrix of $\sqrt{n_g}\, r_g$ is (Neudecker and Wesselman, 1990)

$$\Psi_g = \tilde{D} \Pi_g \Phi_g \Pi'_g \tilde{D}', \tag{50}$$

where

$$\Pi_g = [I - (I \otimes P_g) K_d][(\Sigma_g)_d^{-1/2} \otimes (\Sigma_g)_d^{-1/2}] \tag{51}$$

and Φ_g is the asymptotic variance matrix of s_g described in section above. Here K is the commutation matrix (see Magnus and Neudecker, 1988), and note that $\tilde{D}$ is the duplication matrix for zero-axial symmetry. The expression for $\bar{\Psi}$, the variance matrix common to all groups, is obtained by replacing in the expression Ψ_g the matrices Π_g, Φ_g, P_g and Σ_g by $\bar{\Pi}$, $\bar{\Phi}$, $\bar{P}$ and $\bar{\Sigma}$, respectively.

Consistent estimators of the DF and NT expressions of Ψ_g are

$$A_g \equiv \tilde{D} \hat{\Pi}_g W_g \hat{\Pi}'_g \tilde{D}'$$

and

$$A_g^* \equiv \tilde{D} \hat{\Pi}_g W_g^* \hat{\Pi}'_g \tilde{D}',$$

respectively, where W_g and W_g^* are given in (42) and (45), respectively, and $\hat{\Pi}_g$ is the matrix of (51) with P_g and Σ_g replaced by R_g and S_g, respectively. The corresponding expressions for $\bar{A}$ and $\bar{A}^*$ are obtained by replacing in the expressions above W_g, S_g, and R_g by $\bar{W}$, $\bar{S}$, and $\bar{R}$, respectively. Here $\bar{R}$ is the correlation matrix associated with $\bar{S}$. In contrast with the test of equality of augmented moment matrices discussed in the last section, now under normality (NT) the null hypothesis H_0 does not imply Assumption A.

The DF and NT chi-squared tests statistics of (29) will then be

$$T = nr'(H \otimes \bar{A}^{-1})\, r$$

and

$$T^* = nr'(H \otimes \bar{A}^{*-1})\, r.$$

The number of degrees of freedom of the test is equal to $k=(G-1)p$. We note that to compute T^* the following equality can be useful (Jennrich, 1970; Neudecker and Satorra, 1996).

$$A_g^{*\,-1}=\tfrac{1}{2}\tilde{D}'[R_g^{-1}\otimes R_g^{-1}-2(I\otimes R_g^{-1})\,JU_g^{-1}J'(R_g^{-1}\otimes I)]\,\tilde{D}, \quad (52)$$

where $U_g \equiv I+R_g \times R_g^{-1}$ and J' is the matrix that converts $\operatorname{vec} X$ to $x=X_d 1$ (i.e., $J' \operatorname{vec} X=X_d 1$) for any square matrix X.

The statistic T^* can be scaled to $\bar{T}=T^*/a$, where $a=\operatorname{tr}(\bar{A}^{*\,-1}\bar{A})/p$. When the distribution of the observed variables is elliptical, then Ψ_g is of the form $\Psi_g=(1+\kappa_g)\,\Phi_g^*$, where κ_g is a kurtosis parameter (Neudecker, 1996) and Φ_g^* was given in (44).[5] Further, when $\kappa_g=\bar{\kappa}$, then $\alpha=(1+\bar{\kappa})/p$ and $\bar{T}$ will then be asymptotically an exact chi-squared statistic when H_0 holds.

REFERENCES

Anderson, T. W. (1984). *An Introduction to Multivariate Analysis.* 2nd ed., Wiley, New York.

Andrews, D. W.K. (1987). Asymptotic results for generalized Wald tests. *Econ. Theory* **3** 348–358.

Bentler, P. M. (1989). *EQS Structural Equations Program Manual.* BMPD Statistical Software, Inc., Los Angeles.

Chou, C.-P., Bentler, P. M., and Satorra, A. (1991). Scaled test statistics and robust standard errors for nonnormal data in covariance structure analysis: A Monte Carlo study. *British J. Math. Statist. Psych.* **44** 347–357.

Dik, J. J., and Gunst, M. C. M. (1985). The distribution of general quadratic forms in normal variables. *Statist. Neerlandica* **39** 14–25.

Foutz, R. V., and Srivastava, R. C. (1977). The performance of the likelihood ratio test when the model is incorrect. *Ann. Statist.* **5** 1183–1194.

Hedges, L. V., and Olkin, I. (1985). *Statistical Methods for Meta-Analysis.* Academic Press, New York.

Jöreskog, K., and Sörbom, D. (1989). *Lisrel 7 A Guide to the Program and Applications.* 2nd ed., SPSS Inc., Chicago.

Jennrich, R. I. (1970). An asymptotic χ^2 test for the equality of two correlation matrices. *J. Amer. Statist. Assoc.* **65** 904–912.

Korin, B. P. (1968). On the distribution of a statistic used for testing a covariance matrix. *Biometrika* **55** 171–178.

Magnus, J. R., and Neudecker, H. (1986). Symmetry, 0–1 Matrices and Jacobians, A review. *Econom. Theory* **2** 157–190.

Magnus, J. R., and Neudecker, H. (1989). *Matrix differential Calculus with Applications in Statistics and Econometrics.* Wiley, New York.

Modarres, R., and Jernigan, R. W. (1992). Testing the equality of correlation matrices. *Commun. Statist. Theory Methods* **21** 2107–2125.

[5] We recall that in the case of an elliptical distribution

$$\Phi_g=2(1+\kappa_g)\,D^+(\Sigma_g\otimes\Sigma_g)\,D^{+\prime}+\kappa_g D^+(\operatorname{vec}\Sigma_g)(\operatorname{vec}\Sigma_g)'D^{+\prime}. \quad (53)$$

Modarres, R., and Jernigan, R. W. (1993). A robust test for comparing correlations. *J. Statist. Comput. Simul.* **46** 169–181.

Moore, D. S. (1977). Generalized inverses, Wald's method, and the construction of chi-squared tests of fit. *J. Amer. Statist. Assoc.* **72** 131–137.

Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley, New York.

Neudecker, H. (1994). Distributions of quadratic forms in normally distributed variables. Tempus Lectures Notes. [University of Amsterdam]

Neudecker, H. (1995a). Mathematical properties of the variance of the multinomial distribution. *J. Math. Anal. Appl.* **189** 757–762.

Neudecker, H. (1996). The asymptotic variance matrices of the sample correlation matrix in elliptical and normal situations and their proportionality. *Linear Algebra Appl.* **237/238** 127–132.

Neudecker, H., and Satorra, A. (1996). The algebraic equality of two asymptotic tests for the hypothesis that a normal distribution has a specified correlation matrix. *Statistics & Probability Letters* **30** 99–103.

Neudecker, H., and Wesselman, A. M. (1990). The asymptotic variance matrix of the sample correlation matrix. *Linear Algebra Appl.* **127** 589–599.

Rao, J. N. K., and Scott, A. J. (1984). On chi-squared tests for multi-way contingency tables with cell proportions estimated from survey data. *Ann. Statist.* **12** 46–60.

Satorra, A., and Bentler P M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In *Latent Variables Analysis: Applications for Developmental Research* (A. von Eye and C. C. Clogg, Eds.), pp. 399–419. Thousand Oaks, CA: Sage.

Searle, S. R. (1982). *Matrix Algebra Useful for Statistics*. Wiley, New York.

Wilson, J. R., and Koehler, K. J. (1991). Hierarchical models for cross-classified overdispersed multinomial data. *J. Bus. Econ. Statist.* **9** 103–110.

Wilson, J. R., and Reiser, M. (1993). Transforming hypotheses for test of homogeneity with survey data. *J. Official Statist.* **9** 815–823.

Zhang, J., and Boos, D. D. (1992). Bootstrap critical values for testing homigeneity of covariance matrices. *J. Amer. Statist. Assoc.* **87** 425–428.